

H. VARGA GYULA

IGEKÖTŐS ALAKULATAINK SZÓFAJGYAKORISÁGI JELLEMZŐI

Egy szöveg lényeges tulajdonságaira derülhet fény azáltal, hogy kimutatjuk: az egyes szófajok milyen mértékben uralják az adott szöveget. A kvantitatív kutatási módszerek térhódításával az utóbbi időben örövendetesen megnőtt a szófajstatisztikai vizsgálatok száma. Az ilyen jellegű munkák általában korszerű gépi feldolgozások eredményei.

E módszernek is vannak azonban nehézségei. Az egyik ilyen megoldhatatlannak tűnő probléma: a homonimitás. (A gép ti. nem tudja elkülöníteni a vár igét a vár főnévtől.) Hasonló fejtörést vagy bosszúságot okoznak az igekötős szavak. Közülük ugyanis formálisan csak a szövegszóként előfordulók különíthetők el, így a szófajstatisztikák csak az elváló igekötőket regisztrálják. Az adatok hitelessége így több helyen vitatható, arról nem is beszélve, hogy az igekötőkről aligha kapunk használható információkat.* A hiba nem is a válaszban, hanem a kérdésben bújlik meg: nem az igekötőt, hanem az igekötős szóelőfordulásokat kell vizsgálni. Így járunk el az itt következő rövid szófajstatisztikai jellemzésben.

Arra a kérdésre keressük a választ, hogy a különböző műfajú szövegekre mennyire jellemző az igekötős alakulatok (tkp. szövegszók) jelenléte, s milyen szófaji eloszlásban. Másfelől nézve: hogyan, milyen súllyal

* A gyakorisági listák így torzulnak el: pl. egy szövegben az ad 120-szor, a megad 40-szer fordul elő. De az utóbbiból 15 elváló igekötő miatt ad lett, a valódi előfordulások tehát: ad: 105, megad: 55. Nyilván így van ez a többi igével, igenévvél, ill. igekötővel is. (Némileg hasonló problémákat GAVRIN is fölvet az angol szóösszetételek kapcsán, vö: GAVRIN 1973, 196 kk.)

(gyakorisággal) vesznek részt a szövegépítésben a különböző szófajú igekötős alakulatok? Tükröződik-e az általános szófajgyakorisági tulajdonságok a szövegnek egy leszűkített (tudniillik az igekötős szóelőfordulásokra korlátozott) területére? Mindebből kiderül, hogy az egyes szövegfajtákban milyen valószínűséggel s milyen szófaji értékben jelennek meg az igekötős szövegszók.

A kérdések megválaszolásához statisztikai mutatókra van szükségünk. Ezek rávilágítanak az igekötők és az igekötős szavak tipikus viselkedésére.

Az igekötőkről az első jelentősebb statisztikai kimutatást J. SOLTÉSZ KATALIN készítette. A hat évi igekötőről szóló monográfiájában (SOLTÉSZ 1959) az egyes igekötőket tárgyaló fejezetek végén, mintegy függelékként tekintélyes anyagra támaszkodó mutatókat közöl. Tájékoztat az egyes igekötők és főbb jelentései előfordulási arányairól. Mivel sem az egyes szövegek terjedelméről, sem a vizsgált hat elem abszolút előfordulási számáról nem esik szó (csak %-os adatok szerepelnek), így ezeket itt nem tudjuk használni.

JAKAB ISTVÁN tanulmánya ilyen szempontból nem hagy kívánnivalót maga után. A jelentős terjedelmű (4, egyenként százezer belűjelnyi) szövegekből nyert adatok korszerű módon, jól áttekinthető csoportosításban és formában kerültek feldolgozásra. (A feldolgozott szövegekre vonatkozó adatokat l. JAKAB 1976, 13.)

Az ilyen jellegű vizsgálatok ma már viszonylag egyszerűen elvégezhetők. Itt ugyanis nagy segítséget nyújtanak a számítógépek. Az alábbi mennyiségi áttekintés alapját a Kossuth Lajos Tudományegyetem Számítóközpontjában készült konkordencialisták adják. E listákról részletesen ír PAPP FERENC (1968 és 1975).

A statisztikai adatok a következő szövegek alapján álltak össze:

A) Vers: Ady Endre Vér és Arany, A magunk szerelme, A halottak élén, Az utolsó hajók c. kötetei (40 370: 1730)¹

B) Publicisztika: válogatás az Élet és Irodalom 1980. évfolyamának cikkeiből (20 023: 1676);

¹Az első számjegy az összes szövegszó, a második az igekötős szövegszók előfordulási számát jelöli.

C) Tudományos próza: N. P. Buszlenko: Bonyolult rendszerek stimulációja. Műszaki K., Bp. 1972 (Ford.: Tamkó József), 100--120. l. (18 947: 2346);

D) Spontán jellegű írásbeli megnyilatkozás (fogalmazás)¹ (33 600; 3420);

E) Spontán jellegű szóbeli megnyilatkozás (beszélgetés)¹ (28 010; 1578).

(A korpuszok terjedelme eléggé eltérő, ezért elsősorban nem egymáshoz, hanem az átlaghoz kell viszonyítanunk, illetve a kiugró eltérésekre figyelniük. A szófaji tulajdonságokról (megoszlás) mindenestre jól hasznosítható adatokat kapunk.)

JAKAB ISTVÁN idézett állománystatisztikai vizsgálatában -- az egyes igekötők megterheltségét (gyakoriságát) tartva szem előtt -- kitér az "igei és nem igei utótagú alakulatok" eloszlására is (i. m. 46). A négy mintából nyert adatok igen nagy hasonlóságot mutatnak az általam vizsgált, eltérő műfajú és terjedelmű szövegek összesített mutatóival. Nála a 4507 előfordulás "63,19 %-ában igével, 36,81 %-ában nem igével kapcsolódnak az igekötők" (i. m. 47).

Én jóval nagyobb anyagot (majdnem a kétszeresét) vizsgáltam, így az előfordulások abszolút száma is lényegesen különbözik. Az öt korpusz főbb adatai:

összes szövegszó:	148 950
igekötős alakulat:	10 750
ebből elváló:	2 226

Látjuk, hogy az (elváló) igekötők a szövegszóknek kb. másfél %-át teszik ki. (Az igekötős szerkezetek szórendi viszonyaihoz vö. VARGA 1987, 118.) Ám az összes szövegszó 7,22 %-a tartalmaz igekötőt. Ez utóbbi 67,2 %-a verbum finitumos, 32,8 %-a pedig más szófajú. (A 4 %-os, csekélynek mondható eltérés az említett műfaji és terjedelmi különbségből adódik.) A vers műfajában az arány így alakul:

¹ Ez közel 300 ötödik o. tanuló írásbeli és szóbeli megnyilatkozását tartalmazza. Erről részletesebben l. AAPA. XVII/3. 23.

	Jakabnál	itt
igei	73,99 %	74,62 %
nem igei	26,01 %	25,38 %

Meglepő ez az egyezés, hisz más költő más terjedelmű anyagát vettük alapul.

Ezek után nézzük meg, hogy az egyes mintákból milyen adatokat nyerünk.

Az összesítés nemcsak a szófajok szövegbeli gyakoriságát reprezentálja, hanem az igekötős alakulatok morfológiai aktivitását is. Mindenekelőtt le kell szögeznünk, hogy az igekötős szövegszók szófaji eloszlása távolról sem követi az érintett szófajok átlagos szövegbeli frekvenciáját. Az alapos arányeltolódások elsősorban abból fakadnak, hogy itt kevesebb szófaj osztozik a szóelőfordulásokon. Jelentős eltérések mutatkoznak az egyes műfajok adatai között is.

	A		B		C		D		E		Összesen	
	(Adv)		(ÉS)		(Tud.)		(Fog.)		(Besz.)			
	db	%	db	%	db	%	db	%	db	%	db	%
V	1291	74,62	878	52,39	781	33,29	3073	89,85	1201	72,11	7224	67,2
INF	111	6,42	148	8,83	73	3,11	211	6,17	260	16,48	803	7,47
PART	189	10,92	234	13,96	389	16,58	19	0,89	14	0,89	845	7,86
ADV	44	2,54	45	2,68	18	0,77	35	1,02	22	1,39	164	1,53
N	95	5,49	371	22,14	1085	46,25	82	2,40	81	5,13	1714	15,94
Össze-												
sen	1730	100	1676	100	2346	100	3420	100	1578	100	10750	100
NV												
össz.	344	19,88	427	25,48	480	20,46	265	7,75	296	18,76	1812	16,86
V+NV	1635	94,51	1305	77,86	1261	53,75	3338	97,60	1497	94,87	9036	84,06
NV+N	439	25,38	798	47,61	1565	66,71	347	10,15	377	23,89	3526	32,80

1. táblázat

Igekötös szövegszók előfordulása öt mintában

Rövidítések:	V	=	ige	ADV	=	határozói igenév
	INF	=	főnévi igenév	N	=	névszó
	PART	=	melléknévi igenév	NV	=	igenév

NAGY FERENCnek 10 szövegmintából kapott általános szófajgyakorisági adatai szerint a listán a főnév áll az első helyen (25,18 %), ezt követi az ige (13,66 %). A bennünket érdeklő szófajok közül ezután a melléknév következik (11,05 %). Az igenevek között a melléknévi igenév (2,31 %) gyakoribb, mint a főnévi (1,09 %) és a határozói igenév (0,35 %). (NAGY 1980, 39.)

Ha csak az igekötős szóelőfordulásokat tekintjük, a sorrend némileg, az arányok alapvetően megváltoznak. A közölt táblázat adatai azt mutatják, hogy az igekötős szövegszók zömét az igei lemészetűek (ige + igenév) teszik ki (84,06 %), ezek többsége ige. Feltűnően nagy használatbeli eltérés figyelhető meg az egyes műfajok között. A szépirodalmi szövegekben igen magas az igék aránya (Adynál 74,62 %, Vácínál 73,99; de -- Jakab szerint -- a széppróza még verbálisabb; i. m. 47--48). Mindez igazolja, hogy az igekötő elsősorban és alapvetően az igeállomány gyarapításának eszköze. A szófaji értéket megváltoztató továbbképzés erősen korlátozott.

Az utóbbi megállapításnak ellentmondani látszik a tudományos stílus adata (nem ige: 66,71 %, ebből 46,25 % névszó). Az ellentmondás csak a szövegszók szintjén áll fenn. A szövegben ugyanis bizonyos névszók igen gyakran ismétlődnek (feladat, kimenet, megoldás stb.). A lexémák szintjén mérséklődik az eltérés, a morfológiai szerkezetre, ill. a morfológiai aktivitásra tehát jóval kevésbé jellemző.

Kiugró és az egyes műfajokra jellemző a melléknévi igenevek előfordulási aránya. NAGY FERENC adatai is azt mutatják, hogy ez a nagy ingadozás szövegtani tulajdonság (i. m. 36--7). Emiatt nehéz megjósolni az igenevek arányát is. A főnévi igenév előfordulása nem annyira csapongó. Ezek többnyire elérik az igék gyakoriságának egytizedét (8--12 %). Esetünkben az 11,12 %¹. Az összes igekötős igenévnek az igékhez mért aránya 25,08 %

¹Speciális, leszűkített anyagunkat nyilván nem vethetjük össze a teljes szövegre vonatkozó mutatókkal. Ezért itt most az igenevek előfordulását nem a szöveg, hanem az igeelőfordulások függvényében nézzük. Ugyanezt a névszókkal nem tehetjük meg, mert közülük csak a deverbálisokat vizsgáljuk.

NAGY F.: 27,49 %, SZENDE I.: 22,42 %, másnál jóval kevesebb, vö. NAGY I. h., KESZLER 1983. 173). Igaz tehát az az állítás, hogy az ige-kötös igeből szabadon, morfológiai kötöttségek nélkül képezhető igenév.

A műfaji sajátosságokról tudjuk, hogy a különböző műfajú szövegekben a főnevek és a mellénevek előfordulása ¹ rendkívül eltérő lehet, nagyban hozzájárulván ezzel a szöveg- és a műfaji sajátosságok alakításához (vö. pl. NAGY 1980, 34--37; ZSILKA 1974, 15--18, 67--75; KESZLER 1980, 171 kk.). Esetünkben az arányok természetesen mások, hisz itt csak az (i-ge-kötös) deverbális névszókat vesszük számításba. A műfaji jellegzetességek azonban nemhogy nem változnak: még szembeötlőbbek.

A minta: (Ady-vers):	0,074
B " (ÉS):	0,423
C " (Tud. pr.):	1,389
D " (fogalm.):	0,027
E " (beszélg.):	0,067

A teljes anyag:	0,237
-----------------	-------

A névszó/ige arányok az öt mintában

Igen magas névszói arány jellemzi a C mintát, viszonylag magas a B-beli érték is. ZSILKA TIBOR szerint ez a szöveg bizonyos "explicitását" is mutatja (1974, 72). Ezek a számok a fogalmazás választékosságáról és igényességéről is árulkodnak. A névszókat létrehozó morfológiai műveletek elősegíthetik a gondolat elvontabb, tömörítettebb kifejezését, elsősorban természetesen a cselekményesség háttérbe szorítását. A választékosságra és az explicitásra való törekvést jelzi a melléknévi igenév gyakori használata is.

¹ A "névszó" jelzetű csoport (főnév + melléknév) további bontása a mellénevek kis száma miatt nem látszik szükségesnek. Számuk egyébként a publicisztikában a legnagyobb: a 371 névszói értékű szövegszóból 30 a melléknév (12,09 %), a többi mintában jóval 10 % alatt van. Zömük melléknévi igenévből lett alkalmi szófajváltással.

	Nagy F.	Keszler B.	A--E minta
főnévi igenév/ige	0,080	0,094	0,111
melléknévi " /ige	0,169	0,015	0,117
határozói " /ige	0,026	0,017	0,023
Összes igenév/ige	2,828	1,247	2,251

Az igenév/ige arányok

A BUSEMANN-féle koefficiensek (vö. ZSILKA 1974, 71) mintájára kiszámított indexekből látható, hogy az igenevek használatát mutató adataink nem különböznek lényegesen más szerzőknek a teljes korpuszokra érvényes mutatóitól. A táblázat két pontjához szükséges magyarázatot fűzni. A melléknévi igenevek előfordulásának szeszélyes -- jóllehet nem indokolatlan -- ingadozásáról már volt szó. Összesítésünkben igen magas a főnévi igenév indexe. Ezt az E (beszélt nyelvi) minta kiugróan magas előfordulási száma adja.¹ Az ok a megnyilatkozásoknak a beszédszituációhoz kötését, valamint a lezárását szolgáló fordulatok gyakori használata. (Pl. azt szeretném még elmondani, ...; Azt akarom hozzáfűzni, ...; Másnap már haza kellett utaznunk stb.)

Az igekötős alakulatok tehát -- morfológiai tulajdonságaik miatt -- nem követik az általános szófajgyakorisági rendet. Az összes igekötős szövegszó 2/3-a ige, a maradéknak több mint a fele igenév. Ezt a megoszlást keresztezik az általános műfaji jellemzők: a kevésbé verbális műfajokban (főleg a C mintában) megnő a főnevek aránya, az igényes, nagy explicitású szövegekben (B és C minta) általában nagy gyakorisággal szerepelnek a melléknévi igenevek. Igekötős melléknév kevés van, ezek mind melléknévi igenévből lettek szófajváltással.

¹ Ez nemcsak az igekötősekre vonatkozik: az egész E mintában a főnévi igenevek 2,77 %-ot tesznek ki (NAGY F.: 1,09 %; KESZLER B.: 1,64 %).

Irodalom

- AAPA XVIII: (Többen:) A szókincsfejlesztés szükségessége és lehetőségei az általános iskolában. Acta Academiae Paedagogicae Agriensis XVIII/3. (1987).
- GARVIN, Paul L. (1973): A számítógép a természetes nyelv vizsgálatában. In: A nyelvtudomány ma. (Szerk.: Szépe György.) Gondolat K., 1973.
- JAKAB István (1976): A magyar igekötők állományi vizsgálata. NytudÉrt. 91. sz.
- KESZLER Borbála (1983): Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In: Tanulmányok a mai magyar nyelv szövegtana köréből. Tankönyvkiadó, 1983.
- NAGY Ferenc (1980): Kriminálisztikai szövegnyelvészet. Akadémiai K., 1980.
- PAPP Ferenc (1968): Amíg a gépi listából olvasható tanulmány lesz. Nyr. 92: 321--333.
- PAPP Ferenc (1975): Konkordancia: írói szótár előkészítése számítógépen. Nyr. 99: 351--355.
- J. SOLTÉSZ Katalin (1959): Az ősi magyar igekötők. Akadémiai K. 1959.
- VARGA Gyula (1987): Igekötős szerkezeteink szórendi jellemzői. Acta Academiae Paedagogicae Agriensis XVIII/3, 107--119.
- ZSILKA Tibor (1974): Stilisztika és statisztika. Akadémiai K. 1974.

Abstract: (Part of Speech Distribution and Frequency of Words with Verbal Prefixes) Word class statistics, wrongly, only record separable verbal prefixes. We examine here a total of 10,750 words with verbal prefixes found in five texts of different genres (with a sum total of about 150,000 words). The question we want to answer is with what frequency and relative distribution the different classes of words with verbal prefixes (verb, infinitive, participle, gerund and noun) take part in text building.

